Equivalent to JP10-509297

PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: HYBRID WAVEFORM AND MODEL-BASED ENCODING AND DECODING OF IMAGE SIGNALS
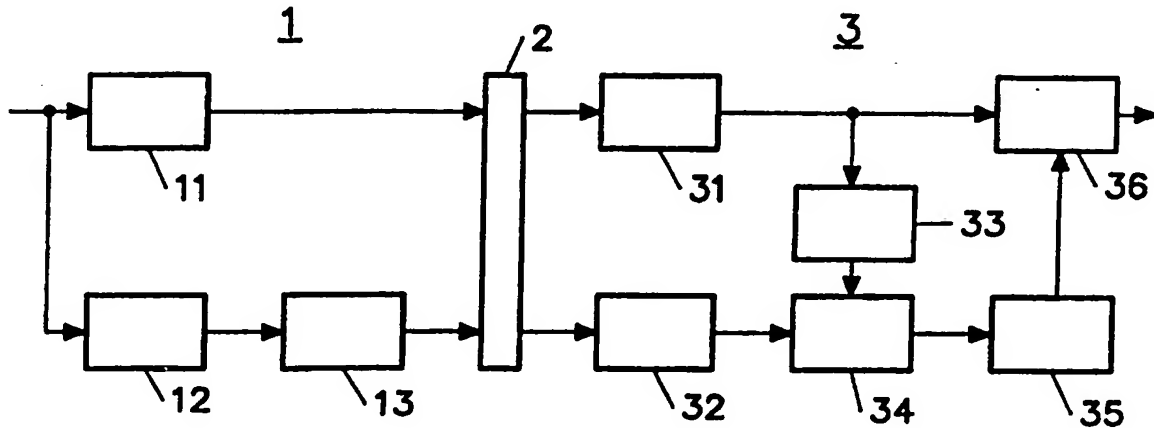
(57) Abstract

A hybrid waveform and model-based image signal encoder and decoder are described. The image signal is waveform-encoded (11) and decoded (31) so as to approximate the waveform of the original input signal as well as possible after compression. In order to compensate its loss, the noise component of the signal (or, more generally, a signal component which gets lost by the waveform encoding) is model-based encoded (12, 13) and separately transmitted or stored. In the decoder, the noise is regenerated (32-34) and added (36) to the waveform-decoded image signal. The invention is particularly applicable to compression of medical X-ray angiographic images where loss of noise leads a cardiologist or radiologist to believe that the image is distorted.

PU030152

CITED

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic	SD	Sudan
CF	Central African Republic		of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

Hybrid waveform and model-based encoding and decoding of image signals.

FIELD OF THE INVENTION

The invention relates to a hybrid waveform and model-based encoder and decoder for encoding and decoding image signals. The invention also relates to a method of transmitting image signals. Transmission is herein also understood to mean the storage of encoded image signals on storage media such as optical or magnetic discs and tapes. The

5   invention is particularly applicable to encoding and compressing medical X-ray angiographic image series.

BACKGROUND OF THE INVENTION

Waveform encoding is a well-known encoding technique used for
10   transmitting of video images. The object of waveform encoding is to approximate the waveform of the original input signal as well as possible at the available bitrate. The quality of the reconstructed signal generally decreases with a decreasing bitrate. Waveform encoders often apply the technique of transform coding. The widely accepted MPEG standard, for example, applies a Discrete Cosine Transform (DCT) to pixel blocks, and quantizes and
15   variable-length codes the coefficients obtained from said transform. Other transforms are currently being investigated. For example, the overlapped transform, e.g. the Lapped Orthogonal Transform (LOT), promises to be suitable for encoding and compressing medical X-ray images. Even at very high compression ratios, the relevant clinical information seems to be present in the compressed images.
20   Model-based coding is a relatively new concept of image data compression. In model-based coding, the original input signal is first modeled, and quantized versions of the model parameters are transmitted to the decoder. Instead of producing a replica of the original waveform, the decoder generates a signal which looks very similar to the original signal, but may have a completely different waveform. A very high compression
25   ratio can be achieved, with model-based coding.

Known hybrid waveform and model-based encoders and decoders are disclosed in "Model-based/waveform hybrid coding for videotelephone images", International Conference on Acoustics, Speech and Signal Processing (ICASSP) 91; Vol.4; 1991; pp.

2741-2744. The known arrangement applies model-based coding to head-and-shoulder images, and waveform encoding to the background image which cannot be modeled. The waveform coding is also used to cancel errors of the model-based analysis system.

5   OBJECT AND SUMMARY OF THE INVENTION

It is an object of the invention to further improve the image quality of waveform-encoded images, in particular medical X-ray images.

According to the invention, a hybrid waveform and model-based encoder is provided for encoding an image signal, comprising a waveform encoder for waveform-

10   encoding the image signal and a model-based encoder for model-based encoding a component of the image signal, at least in as far as said component is absent in the waveform-encoded image signal.

The invention is based on the recognition that, at very high compression ratios, a change of the characteristics in the original images can be observed. This is caused,

15   *inter alia*, by the fact that high-frequency components of the image are very coarsely quantized or even totally removed. The hybrid encoder according to the invention provides an effective and efficient way of regenerating the characteristics of said component.

More particularly, the noise in an image signal is reduced or removed by waveform encoding. This results in an impression of reduced sharpness, which is very

20   disadvantageous in medical applications. Although most clinical information seems to be present in the decompressed X-ray angiographic images, a cardiologist will appreciate a sharp image. As he is accustomed to the presence of noise, the absence thereof leads him to believe that the image is seriously distorted. The transmission of model-based encoded noise requires an extremely low bitrate, and the noise can easily be regenerated in the decoder.

25   Even at very high compression ratios, the original and decoded images are perceptually almost the same.

In an embodiment of the invention, the model-based encoder is arranged to encode the power spectral density of the image signal. The power spectral density of the waveform-decoded signal is determined at the receiver end . Both spectra are then compared

30   and the difference is used to control the regeneration of noise.

In a further embodiment, the waveform-encoded signal is decoded at the transmitter end, and the difference between the original image signal and the waveform-decoded signal is determined. Model-based encoding is then applied to said difference signal.

The image signal is Preferably subjected to transform coding. This is particularly attractive because the same transform circuit is used for both waveform encoding and model-based encoding.

It is to be noted that the above-mentioned publication "Model-based/waveform hybrid coding for videotelephone images" states that the performance of waveform encoders can be improved by the incorporation of model-based coding. However, this aspect is not further elaborated. The particular feature of applying model-based coding to a signal component, which would get lost by waveform encoding and compression, is not disclosed.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1 and 2 show diagrams of a first and a second embodiment of a system comprising a hybrid encoder and decoder according to the invention.

Figs. 3 and 4 show more detailed embodiments of the system shown in Figs. 1 and 2, respectively.

Figs. 5-7 show amplitude spectrums to illustrate the operation of the system shown in Figs. 3 and 4.

Fig.8 shows a graph of a luminance-adaptive scaling factor for improving the system shown in Figs. 3 and 4.

DESCRIPTION OF PREFERRED EMBODIMENTS

Fig.1 shows a diagram of a first embodiment of a system comprising a hybrid encoder and decoder (herein after sometimes abbreviated to "codec") according to the invention. The system comprises a hybrid encoder 1, a transmission or storage medium 2, and a hybrid decoder 3. In the hybrid encoder, the input image signal is applied to a waveform encoder 11 and to a model-based encoder. The model-based encoder includes an estimation circuit 12 for estimating model parameters representing global information about the image signal, e.g. the shape of its power spectral density or the energy of noise components in this signal, and an encoding stage 13. The waveform-encoded signal and the model-based encoded signal are then applied to transmission or storage medium 2. In the hybrid decoder, the waveform-encoded image signal is decoded by a waveform decoder 31, whereas the model-based encoded signal is decoded by a decoding stage 32. The waveform-decoded signal is applied to an estimation circuit 33 which is arranged to estimate the same model parameters as are estimated by circuit 12 of the encoder. In a comparison circuit 34,

the decoded model parameters are compared with the parameters calculated by estimation circuit **33**. The difference thus obtained is an indication of that information which could not be represented well enough by the waveform codec. This information is regenerated in a regeneration stage **35** and added to the waveform-decoded image signal by means of an adder

5   circuit **36**.

Fig.2 shows a diagram of a second embodiment of the codec. The same reference numerals are used for the same circuit elements as shown in Fig.1. The hybrid encoder **1** now comprises a waveform decoder **14**, which locally decodes the waveform-encoded image signal, and a subtracter **15** which subtracts the locally decoded waveform

10   from the original image waveform. The model-based encoding, performed by estimation circuit **12** and encoding stage **13**, is now applied to the difference signal. This embodiment is more efficient in terms of bitrate because model-based encoding is applied only in as far as the relevant component is indeed absent in the waveform-encoded signal. Moreover, the corresponding decoder is less complicated because the estimation circuit and comparison

15   circuit (**33** and **34**, respectively, in Fig.1) can now be dispensed with.

A more detailed example of model-based coding of noise will now be explained with reference to the system shown in Fig.3. In this embodiment, the image is divided into blocks of pixels and the model-based coding of noise is applied to the transform coefficients, obtained by transform coding (e.g. DCT or LOT) said pixel blocks. The system

20   is more or less similar to the embodiment of Fig.1, except that the model parameters are calculated in the transform domain.

In the hybrid encoder **1** of Fig.3, the waveform encoder comprises a transform circuit **101**, a quantizer **102** and a variable-length coder **103**. Such a waveform encoder is generally known and described, for example, in "On Adaptive DCT Coding

25   Techniques for Digital Video Recording", IERE Proc., 7th Int. Conf. Video, Audio & Data Recording, 22-24 March, York (UK), pp.199-204. Hereinafter, a Lapped Orthogonal Transform (LOT) is assumed to be used. The LOT transforms overlapping blocks of 2M*2N pixels into blocks of M*N coefficients. The overlapping pixel blocks are denoted $p_{i,j}$ in which i indicates the vertical position ($i=1..N_v$) and j indicates the horizontal position j ($j=1..N_h$)

30   of the block in the image. The coefficients are denoted $c_{i,j}[m,n]$ ($m=0..M-1$, $n=0..N-1$). For a more complete description of the LOT, reference is made to "The LOT: Transform Coding Without Blocking Effects", IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol.37, No.4, April 1989.

Fig.5 shows a typical amplitude spectrum (square root of the average energy as a function of the spectral frequency) of an original X-ray image. The square values of transform coefficients represent the energy and show a similar distribution. Fig.6 shows the spectrum after waveform encoding and decoding. The loss of most of the high-frequency

5   information can clearly be observed. The system shown in Fig.3 substitutes the lost energy at these frequencies for noise generated at the decoder side.

The model-based encoder in Fig.3 comprises transform circuit **101**, an estimation circuit **104**, and encoding stage **105**. Estimation circuit **104** first calculates, for the current image, the energy of all transform coefficients $c_{i,j}[m,n]$, except the DC coefficient

10   $c_{i,j}[0,0]$, in accordance with.

$$\sigma^2[m,n] = \frac{1}{N_h \cdot N_v} \sum_{i=1}^{N_h} \sum_{j=1}^{N_v} (c_{i,j}[m,n])^2 \qquad\qquad \text{Eq.1}$$

where it has been assumed that the mean value of $c_{i,j}[m,n]$ is equal to zero. The total MN-1 energies $\sigma^2[m,n]$ give an indication of how the energy in the original image is distributed among the different spatial frequencies. The energies $\sigma^2[m,n]$ are lossless coded by encoding

15   stage **105**, by well-known differential coding and variable-length coding techniques. The bitrate required is small as compared with the total number of bits required for the quantized transform coefficients. For images of resolution 512*512 pixels, and for M=N=32, 512^2 quantized transform coefficients have to be transmitted, whereas only 32^2-1 noise energies are needed. The noise energies thus constitute only about 0.4% of the total number of

20   parameters. Furthermore, the block of MN-1 noise energies looks rather smooth, as is illustrated in Fig. 7. They can be zigzag-scanned in a similar way, as is known in the art of transform coding, and then differentially coded.

In the hybrid decoder 3 in Fig.3, the waveform decoder comprises generally known circuit elements such as variable-length decoding stage **301**, inverse

25   quantizer **302**, and inverse transform circuit **303**. The model-based decoder comprises estimation circuit **304**, decoding stage **305**, subtracter **306** and noise generator **307**. The transform coefficients applied to estimation circuit **304** are denoted $\hat{c}_{i,j}[m,n]$.

Estimation circuit **304** calculates the energies $\hat{\sigma}^2[m,n]$ of the reconstructed coefficients $\hat{c}_{i,j}[m,n]$ in a similar way as estimation circuit **104** in the encoder. Thus:

30   $$\hat{\sigma}^2[m,n] = \frac{1}{N_h \cdot N_v} \sum_{i=1}^{N_h} \sum_{j=1}^{N_v} (\hat{c}_{i,j}[m,n])^2 \qquad\qquad \text{Eq.2}$$

Decoding stage 305 decodes the energies $\sigma^2[m,n]$ as transmitted by the model-based encoder. Subtracter 306 calculates the difference

$$\Delta[m,n] = \max(0, \sigma^2[m,n] - \hat{\sigma}^2[m,n])$$ Eq.3

In response to the difference signal thus obtained, noise generation circuit 307 produces
coefficients representing random noise with a variance

$$\sigma_\Delta^2[m,n] = k \cdot \Delta[m,n]$$ Eq.4

in which k is a constant. The noise coefficients are added by an adder 308 to transform coefficients $\hat{c}_{i,j}[m,n]$, for all pixel block positions i and j.

Optionally, the regeneration of noise may be made dependent on the
condition

$$p[m,n] > t[m,n]$$ Eq.5

in which t[m,n] is a predetermined threshold which may depend on the spatial frequencies m and n, and p[m,n] is the percentage of lost energy defined by

$$p[m,n] = \frac{\Delta[m,n]}{\sigma^2[m,n]} \cdot 100$$ Eq.6

The values of k and t[m,n] can be found experimentally, taking the perceptual similarity of the original and the coded images as a criterion. The constant k may even be chosen to be different for different spatial frequencies m and n.

As a result of the addition by adder 308, new coefficients $\tilde{c}_{i,j}[m,n]$ are generated in accordance with

$$\tilde{c}_{i,j}[m,n] = \hat{c}_{i,j}[m,n] + x_{i,j}[m,n]$$ Eq.7

where $x_{i,j}[m,n]$ represents a sample of a random noise generator with variance $\sigma_\Delta^2[m,n]$, which has a probability density function which is equal as much as possible to that of the noise in the original image.

It is to be noted that the calculation of lost energy is performed in an 'asymmetric way': negative values of $\Delta[m,n]$ are rounded off to zero. This is done because an increase of the energies of the LOT coefficients due to their quantization can never be compensated by adding a random noise signal.

The system shown in Fig. 3 can be transformed into a system which is more similar to the one shown in Fig.2. This further embodiment is shown in Fig.4. Estimation circuit **104**, already discussed, calculates the energies $\sigma^2[m,n]$ in accordance with Eq.1. A further estimation circuit **106** calculates the energies $\hat{\sigma}^2[m,n]$ of the coefficients

5    $\hat{\sigma}^2[m,n]$ applied by quantizer **102**, in accordance with Eq.2. The random-noise variances $\sigma_\Delta^2[m,n]$ and percentages p[m,n] are now calculated, in accordance with Eqs. 4-6, by a calculation circuit **107** in the encoder instead of in the decoder (setting $\sigma_\Delta^2[m,n]=0$ if p[m,n] < t[m,n]). The random-noise variances $\sigma_\Delta^2[m,n]$ are lossless coded by encoding stage **105** and transmitted to the decoder. In the decoder, the noise generation is performed by

10   decoding stage **305** and noise generator **307**, in a similar way as in the system of Fig.3.

The noise model assumed hereinbefore for the X-ray images is a very simple one. In reality, the noise may not have stationary statistics within an image. It has been observed, for example, that the noise energy depends on the local luminance in the image. Furthermore, it has been observed that the power spectral density of the noise also

15   depends on the local luminance. It changes from rather flat in low-luminance areas to a more low-pass characteristic in high-luminance areas.

Without noise regeneration, the loss of the noise is most apparent in the low-luminance areas. Simulations have shown that the noise coding method sometimes does not add enough noise in these areas, and too much noise in the bright areas. Of course, for

20   each block of transform coefficients, more detailed information about the shape of the noise spectrum could be transmitted to the decoder. This would, however, significantly increase the bitrate. A fairly simple solution, which does not require an additional bitrate, is to adapt the variance of the noise regenerated in the decoder to the local luminance in the image. As is well-known, the DC coefficient $DC_{i,j}=\hat{c}_{i,j}[0,0]$ represents the mean luminance of the pixel

25   block $p_{i,j}$. Instead of adding noise with a variance $\sigma_\Delta^2[m,n]$, noise is added with variance

$$s(DC_{i,j}) \cdot \sigma_\Delta^2[m,n] \qquad\qquad\qquad\qquad\qquad Eq.8$$

in which $s(DC_{i,j})$ is a scaling factor depending on $DC_{i,j}$. Fig.8 shows an experimentally determined function which has been used to adapt the variance of the noise regenerated in the decoder to the local luminance. In Fig.8, the vertical axis indicates the scaling factor $s(DC_{i,j})$,

30   whereas the horizontal axis indicates the luminance value.

In summary, a hybrid waveform and model-based image signal encoder and decoder are described. The image signal is waveform-encoded so as to approximate the

waveform of the original input signal as well as possible after compression. In order to compensate its loss, the noise component of the signal (or, more generally, a signal component which gets lost by the waveform encoding) is model-based encoded and separately transmitted or stored. In the decoder, the noise is regenerated and added to the waveform-

5   decoded image signal. The invention is particularly applicable to compression of medical X-ray angiographic images, where loss of noise leads a cardiologist or radiologist to believe that the image is distorted.

Claims

1.   A hybrid waveform and model-based encoder for encoding image signals, comprising a waveform encoder for waveform-encoding the image signal and a model-based encoder for model-based encoding a component of the image signal, at least in as far as said component is absent in the waveform-encoded image signal.

5   2.   An encoder as claimed in Claim 1, wherein said component is formed by the noise which is present in said image signal.

3.   An encoder as claimed in Claim 1 or 2, wherein the model-based encoder is arranged to encode the power spectral density of the image signal.

4.   An encoder as claimed in Claim 1 or 2, comprising a waveform decoder

10   for decoding the waveform-encoded image signal, means for determining a difference signal between the image signal and the waveform decoded image signal, said difference signal being applied to the model-based encoder.

5.   An encoder as claimed in Claim 4, wherein the model-based encoder is arranged to encode the variance of the noise in said difference signal.

15   6.   An encoder as claimed in any one of Claims 1-5, comprising a transform arrangement for transforming the image signal into coefficients, wherein the model-based encoder is arranged to apply model-based encoding to said coefficients.

7.   A hybrid waveform and model-based decoder for decoding image signals, comprising a waveform decoder for decoding a waveform-encoded version of the image

20   signal, a model-based decoder for decoding a model-based encoded version of the image signal, means for determining to what extent a component of the image signal is absent in the waveform-decoded image signal, and means for adding said missing component to the waveform-decoded image signal.

8.   A hybrid waveform and model-based decoder for decoding image signals,

25   comprising a waveform decoder for decoding a waveform-encoded image signal, a model-based decoder for decoding a model-based encoded component which is absent in the waveform-encoded image signal, and means for adding said decoded component to the waveform-decoded image signal.

9.   An decoder as claimed in Claim 7 or 8, wherein said component is representative of the noise which is absent in said waveform-encoded signal.

10.   An decoder as claimed in Claim 7, 8 or 9, wherein the model-based decoder is adapted to decode the power spectral density of the model-based encoded image signal.

11.   A method of transmitting image signals, comprising the steps of waveform-encoding the image signal, model-based encoding a component of the image signal in as far as said component is absent in the waveform-encoded image signal, transmitting said waveform-encoded image signal and said model-based encoded component, decoding said waveform-encoded image signal and model-based encoded component, and combining said decoded image signal and model-based decoded component.

12.   A method as claimed in Claim 11, wherein said predetermined component is formed by the noise which is present in said image signal.

13.   A method as claimed in Claim 11 or 12, wherein the step of model-based encoding comprises the step of encoding the power spectral density of the image signal.

14.   A method as claimed in Claim 11 or 12, comprising the step of waveform-decoding the waveform-encoded image signal, determining a difference signal between the image signal and the waveform decoded image signal, and model-based encoding said difference signal.

15.   A method as claimed in Claim 14, wherein the step of model-based encoding is applied to the variance of the noise in said difference signal.

16.   A method as claimed in any one of Claims 11-15, comprising the step of transforming the image signal into coefficients, the model-based encoding being applied to said coefficients.

17.   An image signal, comprising a waveform-encoded image signal and a model-based encoded signal representative of at least a predetermined component of the image signal in as far as said component is absent in the waveform-encoded image signal.

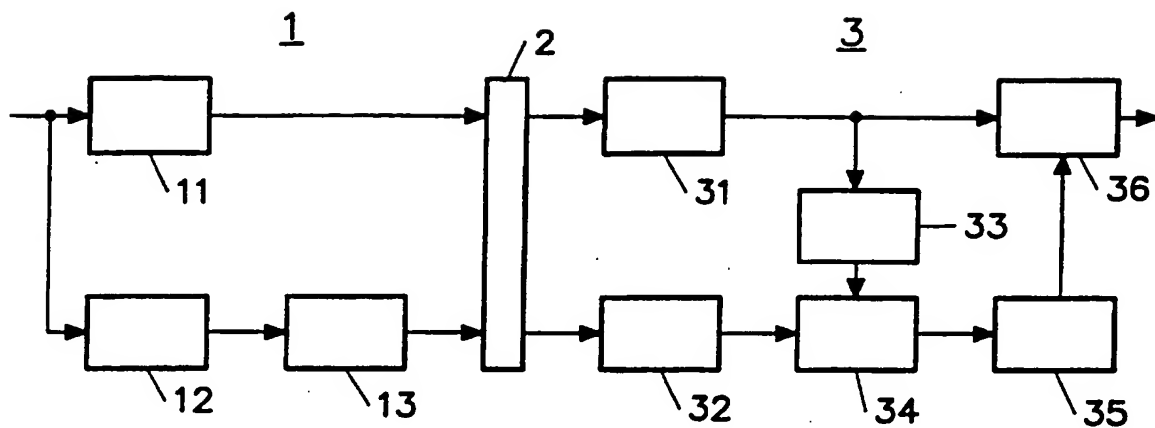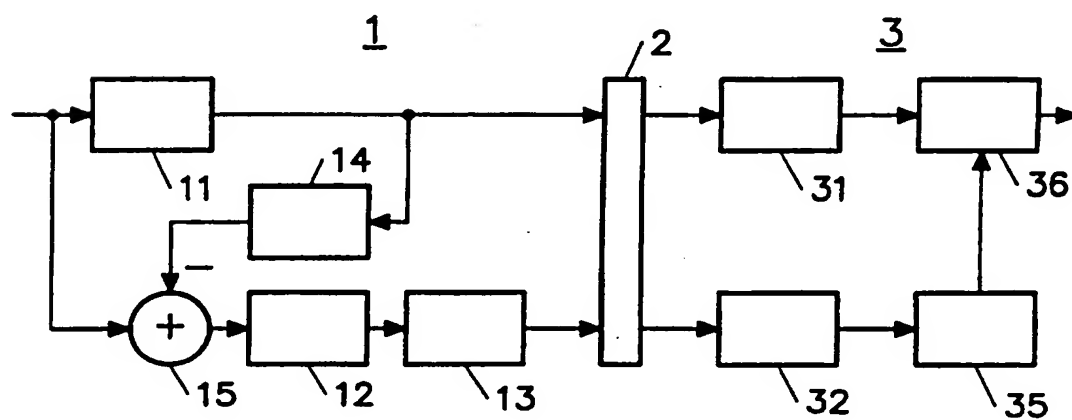18.   A storage medium on which an image signal as claimed in Claim 17 has been stored.
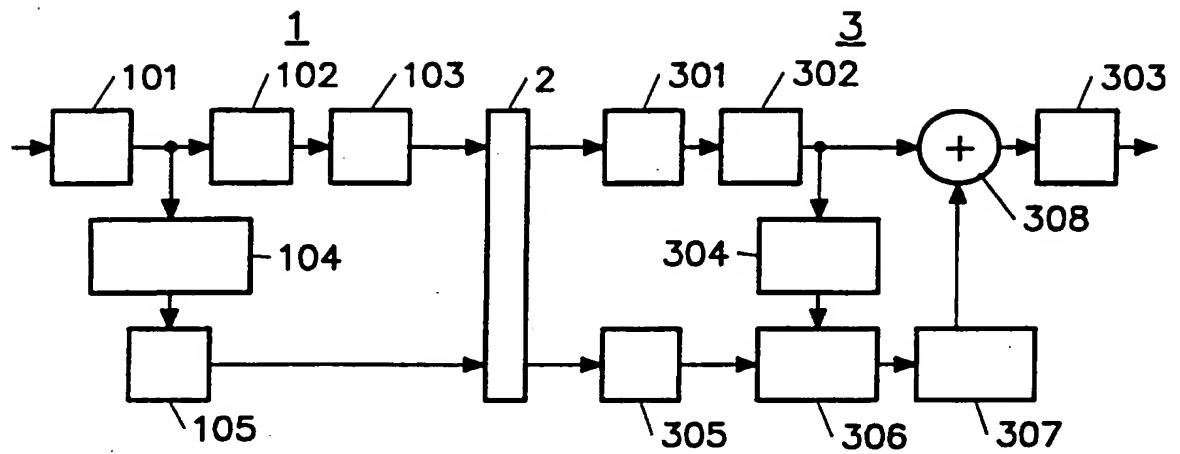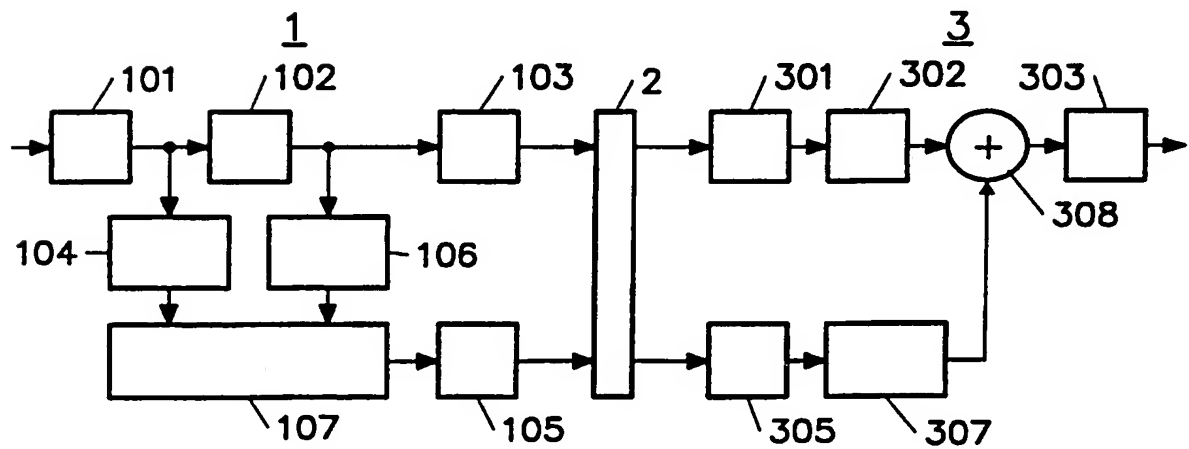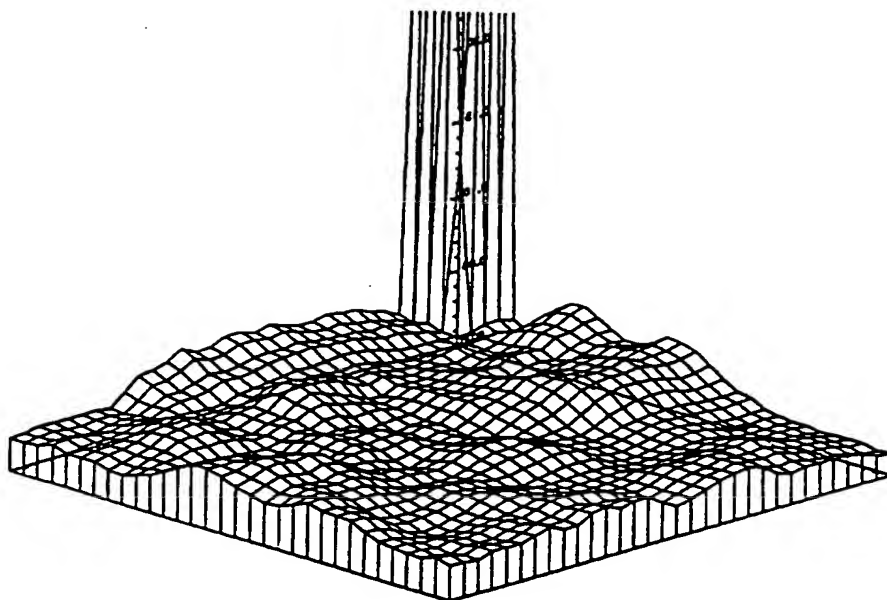
FIG. 1

FIG. 2

FIG. 3
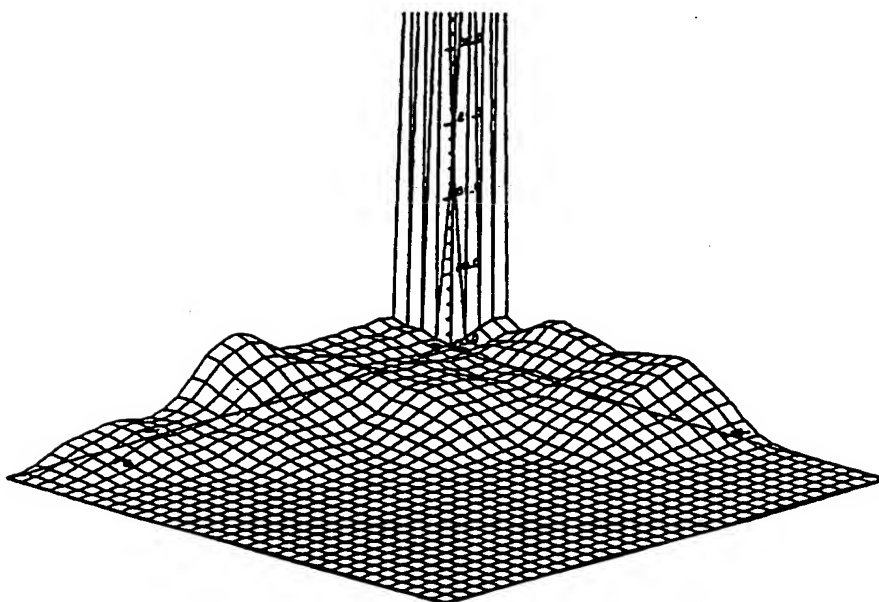

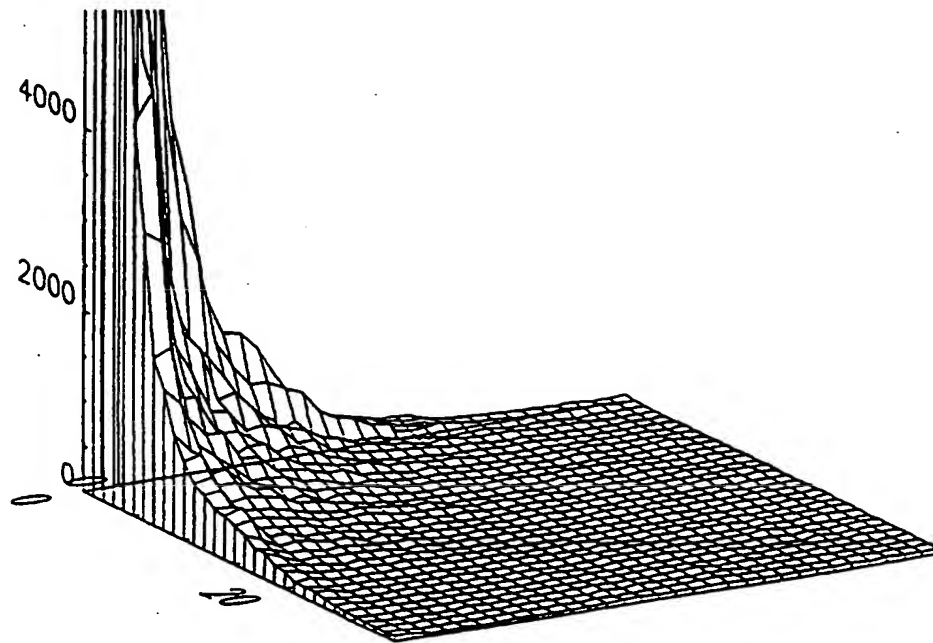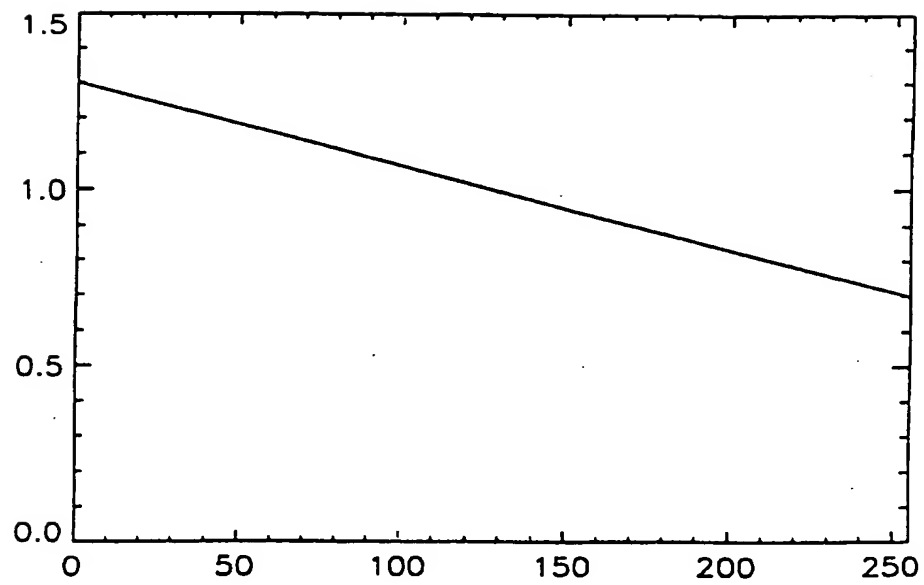
FIG. 4

FIG.5

FIG.6

4/4

FIG. 7

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB 96/00880

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: H04N 7/26

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	IEICE Transactions on Communications, Volume E75-B, No 5, May 1992, (Tokyo, Japan), Y. Nakaya et al, "Model-Based/Waveform Hybrid Coding for Low-Rate Transmission of Facial Images" page 377 - page 384	1,4,7,8,11, 14,17,18
A		2,3,5,6,9, 10,12,13,15, 16
	--	
X	EP 0614318 A2 (KABUSHIKI KAISHA TOSHIBA), 7 Sept 1994 (07.09.94), page 7, line 25 - line 36	1,4,7,8,11, 14,17,18,
A		2,3,5,6,9, 10,12,13,15, 16
	--	

X	Further documents are listed in the continuation of Box C.	X	See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" erlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 February 1997	2 1 -02- 1997
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM	Authorized officer Anders Ströbeck